

The Concept of Emergence in Complexity Science: Finding Coherence between Theory and Practice

Markus Christen* and Laura Rebecca Franklin**

*Institute of Neuroinformatics, University & ETH Zurich

**Department of Philosophy, Columbia University, New York

The concept of *emergence* plays a central role in the science(s) of complexity. Yet while its *importance* seems substantial, its *clarity* is not. Some researchers provide only vague definitions of emergence, or define it by example rather than by providing necessary and sufficient conditions of application. Furthermore, a plurality of discordant emergence concepts can be found in the literature. In this paper we construct a philosophical framework in which we consider candidate emergence concepts, later classifying them according to explanatory power. After clarifying which types of emergence are both coherent and interesting, we apply the framework to paradigmatic examples of 'emergence' in complex systems research. We also provide survey data about what SFI researchers and students consider *emergence* to be. We conclude that the emergence of interest to complexity science researchers is not the causally 'interesting' form of emergence which philosophers discuss; 'emergence' in complexity science, while intriguing to many, is still fully consistent with most forms of reductionism and does not allow for downward causation.

1 the importance of being emergent

Many phenomena that arise out of complex, adaptive systems are called 'emergent' by complexity science researchers (see Waldrop (1992), Holland (1998)). Such emergent phenomena are considered to be *unpredictable* and *irreducible*. Some claim that emergent entities also have the property of *macrodetermination* or *downward causation*; as when an emergent entity causally influences entities at a lower level (i.e. a level that can only be described using a more fine-grained representation). Kim (1999) states that emergent entities *must* exhibit the property of downward causation to be of scientific interest. By emphasising the importance of emergence, the science of complexity positions itself contrary to both *reductionism* and *vitalism*, theses which will be defined as the discussion unfolds. Emergence crops-up in a variety of fields; in the life-sciences, *life* itself is claimed as emergent; mind-brain-interaction also seems home to emergents, with consciousness and qualia being prominent candidates; social scientists call on emergence when characterising everything from stock-market fluctuations to voting outcomes. In sum – emergence is everywhere...but what *is* it?

To answer this question, we need to clarify the different possible ways of using the concept of emergence (Hoyningen (1994)), the ontological status of emergent phenomena, and the role of the observer in emergent systems (Casti 1986). Some researchers have attempted to formalise the notion of emergence (see, for example Baas (1994) and Shalizi (2001)). A further task is to see whether these formalised emergence concepts, in addition to some which we will construct, match up with the concept as it is usually used by complexity scientists. A philosophical framework will aid us. In section 2 we explain the assumptions of all versions of emergentism. A specific framework is developed in section 3 which allows us to address the following questions: 1) what is called emergent? 2) which epistemic or nomological claims are associated with these emergent entities? 3) what is the status of emergent entities in scientific theory? 4) is emergence compatible with, or even reliant on, a reductionist approach? In section 4 we consider which concepts of emergence are interesting, and in section 5 we address whether these 'interesting' concepts are actually the ones used by complexity researchers. To do this, we analyse emergence case studies as well as the results of a survey of SFI researchers vis-à-vis emergence.

2 basic assumptions of the emergence concept

The concept of emergence is usually put to use in the context of two metaphysical claims: 1) ontological monism—there is just one kind of stuff in the world; 2) hierarchical realism—in the world there exist multiple legitimate levels of complexity and organisation. In this section we will define these assumptions, and contrast emergentism, a.k.a. non-reductive materialism, to two rival positions, reductionism and vitalism.

2.1 assumption one: ontological monism The first of assumption of emergentism, *ontological monism*, has its roots in the writings of the British Emergentists of the 1920s. Emergentism as a philosophical position was formulated by C.D. Broad and others as a compromise between

reductionism (called "mechanism" at that time) and *vitalism* (Broad (1925)). The vitalist position claimed the existence of a life-substance, inherently different from the inanimate stuff found in rocks and clouds, which accounted for life's unique properties. Promoted by Henri Bergson, this position seemed suspect to many scientists, as it postulated entities (like Bergson's 'elan vitale') which were out of the reach of scientific investigation. Seeming to be unparsimonious and unscientific, emergentists and reductionists ruled out vitalism and embraced ontological monism.

2.2 assumption two: hierarchical realism The second assumption of emergentism is that the system¹ under consideration can be divided into hierarchical levels. There must be at least two levels in a given system — a 'lower' level, which we'll call 'A', consisting of the parts, and an 'upper' level, called 'B', consisting of the whole system. The two levels may be connected in the following two ways — *microdeterminism* claims that the parts of the system and their interactions on level A, the lower level, fully determine the behaviour of the entire system. On the other hand, if level B acts causally on level A, we call this *macrodetermination* or *downward-causation*, a possible trademark of emergence. Note that the decomposition of a level into parts, as well as the whole system into levels, are not unproblematic tasks and may depend on the objectives of a particular investigation. Hierarchical realism claims only that the system is partitionable into levels, not that those levels must always be independent of the pragmatics of investigation.

It is a legitimate question whether or not the world *really is* ordered into hierarchical levels or if this hierarchy is only an artefact of our way of seeing and thinking about the world. While not discussing this question at length, we want to mention an argument of Herbert Simon (Simon (1962)), claiming that systems in the world of evolutionary origin are likely to be hierarchically structured. This hierarchy results from the error-robust construction principle, which he clarifies with the tale of the two watchmakers; Hora and Tempus both produced watches consisting of 1000 parts each. Hora produces his watches in a modular way, with 10 parts making up a single module, 10 modules used in a bigger module, and ten of those finally going into a watch. On the other hand, Tempus builds his watches in a single production process. When Hora is disturbed, he only has to restart the construction of a single module, while Tempus has to restart the construction of the whole watch; his process is therefore much less robust and he needs more time than Hora to finish a watch. Because of the increased robustness of Hora's construction method, Simon argues that entities of the world, which evolved under disruptive conditions, are likewise likely to be organised hierarchically.

2.3. emergence and reduction: The concept of emergence – at least in the philosophical tradition – is opposed to reductionism. The enthronement of non-reductive materialism and the fading away of classical reductionism since the 1970s has contributed to the new boom of "emergence", as noted by Kim (Kim (1999)). Unlike the vitalist view-point, the reductionist position is not disputed on ontological grounds (Stöckler (1991)) – its ontology is perfectly sparse and only committed to whatever entities physics countenances; rather, emergentists challenge the epistemic, methodological and explanatory claims of reductionism (see for example Beckermann et al. (1992)). The contrast between emergentism and reductionism can be understood as follows—consider again two hierarchical levels, A and B, formed by different classes of entities, e.g. molecules on level A, and cells on level B. For each level we have a theory that defines the entities on that level, their organisation in terms of structure or function, and explains their behaviours. Emergentists claim that, despite the fact that the entities of the higher level must be *consistent with* the entities and theories at the lower level, level B need not be *reducible* to level A.

But just what does 'reducible' mean? There are a few reduction concepts to be considered. First, with the assumption that scientific knowledge is axiomatizable², *epistemic reductionism* claims that knowledge about level B can be *deduced* from knowledge of level A (see Nagel (1961) and Oppenheim et al. (1958). Kim (1999) offers a different, functional reduction procedure). *Methodological reductionism* claims that the methods of level A should be used to explain phenomena

¹ There is no common understanding concerning the concept "system". Under "system", we understand any entity which is in a certain kind separable from its environment and is object of a scientific study.

² Many scientists think that at least some parts of the world can be axiomatized. John Holland mentions that it is possible that "the parts of the universe that we can understand in a scientific sense – the parts of the universe that we can describe via laws (axioms, equations)—constitute a small fragment of the whole" (Holland (1998), p. 231). However, it is that small fragment which is the concern of organized science and reductionists.

on level B (strong version) or that *post facto* we will eventually see how the methods of level A can be applied to level B. Finally, *explanatory reductionism* claims that for each phenomena on level B there exist a mechanism on level A which explains the phenomena. The theory of emergentism has been developed to oppose at least one of these three concepts of reductionism, while not falling into vitalism by postulating a new type of substance altogether. In section 3 we will go into more detail about each of these reduction concepts, and explain which ones emergentists must oppose if their position is to be coherent and interesting.

The concept of emergence also opposes epiphenomenalism, which claims that there might be phenomena on the systems level lacking any causal or explanatory power. In this view, emergent phenomena do not effect the lower level in ways that could not be explained by other lower-level phenomena—therefore, one might as well just ignore these epiphenomenal emergents altogether – after all, if they can do neither causal nor explanatory work, what *are* they good for? For example, some fear that our conscious experiences are epiphenomenal, with experience being completely unnecessary in the explanation of human action. That is, it could be that the process of the neuronal inputs and outputs, and the brain’s wiring, would be sufficient to explain human behaviour, and positing intentional states and experiences would add *nothing more* to that explanation. Emergentists reject this. They maintain either of the following positions: 1) the lower level theories, like a purely neurological one, miss some explanations, and possible causations, and on the higher-level some ‘epiphenomenal’ phenomena are actually explanatory and possibly causally efficacious, 2) the lower level might always provide *some sort* of explanation, but high level processes might also provide useful explanations which can be *preferred* for a variety of reasons, including pragmatic ones. Consistent with the first response, those emergentists who support downward causation have an easy time avoiding epiphenomenalism – since emergent entities are supposed to be causally efficacious, and explanations usually involve providing causal explanations, downward causation of emergents buries any concerns one might have about epiphenomenal entities. However, even those holding the second view, that higher level properties provide another explanation for things still explainable on the lower level, deny epiphenomenalism by claiming that in some circumstances these higher level explanations are better, and therefore the higher-level entities are, if not causally efficacious, at least extremely useful in explanations. One could imagine the psychological states might fall into this category – perhaps I could explain my boyfriend’s affectionate behaviour in terms of a neurophysiological state, but isn’t his love for me a more satisfying explanation?

To summarise the developments of this section, consider the chart below, which specifies the stances of the three positions we have considered on ontology, levels, and epiphenomenalism:

↓ theory; assumption ⇒	ontology	<i>independent</i> levels	epiphenomenal entities
reductionism	monist	no	yes
emergentism	monist	yes	not primarily
vitalism	dualist (or more)	yes	not primarily

table 1: theory assumptions

3 what is emergent? the development of a framework

The term ‘emergent’ has been applied to various kinds of entities. We use the term “entity” here as a general term for *whatever* might be called “emergent”. In this section, first we discuss the ontological question: to which kinds of entities can and should the term ‘emergence’ apply? Later, we consider emergence from an epistemic perspective, that is, what is the role of emergence in scientific theory and knowledge? Many terms and concepts will be presented in this section, with some elements remaining unconnected. This section should be seen as a toolbox of concepts which will be called upon in the later sections of the papers.

The table below shows the ontological and epistemic categories we will consider.

ontological categories	epistemic categories
1. properties or behaviours	1. phenomenological level
2. laws	2. level of theory

3. organisation of parts (patterns): macroproperties	
---	--

table 2: framework categories

Note that there is some interplay between the two epistemic levels: emergence on the phenomenological level *may* reflect non-perceptual emergents – but not necessarily. That is, while at first sight some properties seem emergent because we are surprised that they appear, a closer look would not justify their theoretical classification as "emergent".

3.1 ontology: properties The most obvious entities to which 'emergence' could apply would be to *properties*. Properties are characteristics of entities, like 'being red', or 'being soluble'. When we consider whether 'emergence' can be applied to properties, we must further specify to which parts of a system the emergent property affixes; for example, emergent properties can concern either the whole system or to the parts of the system, as in the case of the flocking behaviour of animal groups – should 'emergent' apply to the whole group, or to each animal in the group? Whether emergent properties make sense may hang on the question of *microdetermination*. Recall that microdetermination assumes that the whole system is completely defined by the state of the parts of the system. Consider the case of a computer screen – the image on the screen seems to be microdetermined by the color values of each pixel on the screen – once you specify the pixel values, you've specified the image on the screen as a whole. However, note that microdeterminism does not imply predictability, as the example of chaotic dynamics shows; although a weather system might be fully determined by the states of all particles in the system, sensitive dependence on initial conditions and our inability to measure the system with absolute precision might still prevent prediction of changes in the system over time. Furthermore, microdetermination does not imply determinism of the whole system, as in certain quantum systems; the micro-level can both be irreducibly probabilistic, yet still fully specify the behaviour of the system as a whole. Finally, an option discussed in depth in section 4, there could be no microdetermination, meaning that a single macrostate is compatible with many microstates – this would occur if the microlevel were not causally closed.

3.2 ontology: laws: Laws are another possible entity that might be called 'emergent'. Laws are spatio-temporally unrestricted regularities, which are emergent if the scope of the law is systems-level, and does not make sense on the level of parts (e.g. temperature in thermodynamics). There are two ways an emergent law might relate to lower-level entities. First, the law might pick out a genuine regularity, yet be reducible to the laws of the lower level. On the other hand, emergent laws might be new and irreducible to the lower level. The base to which a reducible-yet-emergent law can be translated might vary, being either functional (based on the causal organisation between parts, irrespective of the particular part identities) or structural (based on the details of the physical elements themselves).

3.3 ontology: organisation: A final entity-type related to emergence is pattern/organisation. The organisation of parts includes their position in space (structure, patterns) and their possible interactions (functional organisation). Some might say that *organisation* would be better included in the aforementioned class, *properties*. However, we separate these ontological categories because the science(s) of complexity stress the importance of patterns in particular. After all, our goal, rather than a desert ontology fitting to the environment in which the writers find themselves, is an ontological framework permitting the characterisation of the candidate emergence concepts in the most natural way possible.

3.4 epistemology: phenomena Science is characterised by more than just theories. Before models can take shape, humans experience phenomena in the world by perceiving stimuli through their senses—we see, hear, taste, feel, and contemplate the entities around us. We've described how emergent entities might be characterised by a deductive relationship, or lack thereof, between entities describable at lower and higher levels of magnification. Another possibility is that we could detect emergents by their phenomenology, that is, by the perceptual experience which they evoke. We are often surprised or shocked by emergents, and perceive that there is something new or different going on with emergents as opposed to 'normal' phenomena. As a graphic example, Chris Langton, pioneer of research into artificial life, claims that when he watched Conway's Game of Life unfolding on his computer screen he, "got then sense that [he] wasn't alone...A completely visceral feeling, hairs standing up on the back of my neck.....I suddenly realised that something on the screen must have

triggered that feeling" (in: Lewin (1999))). While the sensations evoked by emergence might have some value, particularly to direct more rigorous and objective research in new directions, caution should be taken. The perception of emergence is actually a psychological issue, the object of a theory of perception. Unless we want to run the risk that our category of emergence is idiosyncratically dependent on the details of human sensory apparatus, a scientific definition of emergence shouldn't be reliant on emergence phenomenology alone.

More simply, sensations themselves might be considered emergent. For example, the smell of ammonia (NH_3) is Mill's classic example of an emergent property, which he claims cannot be deduced from the smell of its components, H_2 and N_2 (see McLaughlin (1992)). However, Mill's claim is confused. Smell is a *secondary property* of NH_3 , --that is, a property resulting not just from the intrinsic properties of ammonia, but from the interaction of the ammonia with the observer. It is the primary properties of ammonia, that is, the arrangement of the NH_3 atoms in space, which allow for interaction with certain receptor molecules in cells of the nose of the observer, which is reducible to the hydrogen and nitrogen component. Mill-style emergence hence does not seem to be a major concern.

3.5 epistemology: theories Theories order our knowledge about the world. In line with our assumption of hierarchical realism, we expect different theories to apply at different levels of organisation. In the development of modern science, this is indeed the case (e.g. biological theory, chemical theory, physical theory). However, while it is clear that we have theories applicable to different levels, the connection between various theoretical levels remains a matter of debate. In particular, although reductionists claim that – in principle – higher level theories can be reduced to lower level ones, this remains to be actually done³. Note that the 'in principle' argument is impossible to falsify - reductionists claim that theories at different levels can be connected, but no number of failures to connect them could convince them otherwise; technical problems can always be used to explain failures. Influential in the philosophy of science, Nagel (Nagel 1961) claims that the theoretical entities on one level can—in principle--be connected to another level via a translation of the predicates from one theory to another and by using translation rules to connect the levels. While acknowledging higher-level theories, he sees them as piggy-backing on the lower level. Nagel writes as a microdeterminist, thinking that only the lower-level description is key to a complete science. However, as mentioned, there is another possibility: *macrodeterminism*. Macrodeterminism can make one of two claims: weak macrodeterminism claims that there are higher level laws that fill in the gaps left by lower level laws, which are incomplete. Strong macrodeterminism claims that not only can higher level theoretical laws fill in the gaps left by lower level underdeterminism, but in fact the conditions specified by the higher level laws can lead to the *violation* of lower level laws. Note that the emergence character of an entity might be relative to a certain theory: superfluidity of helium might be called 'emergent' in classical mechanics, but, as it is explained in quantum mechanics it might therefore not be called emergent. Emergence which is not relative to a certain theory is called absolute emergence. We will apply these ideas about reduction and the theory-relativity of emergence to the different concepts of emergence mentioned in the next section.

Another epistemological element of relevance is *predictability*. We will consider here two types of predictability—inductive and theoretical. As described by Kim, inductive predictability is the ability, after having observed an emergent entity occurring simultaneous with a particular base, of predicting the emergent entity the next time that base occurs. For example, if we consider pain to be an emergent property, and we note that every time the c-fibers fire, pain results, we might find the experience of pain to be inductively predictable on the basis of c-fiber firing. Inductive predictability doesn't require a full understanding of the system, or of how the base leads to the emergent. Theoretical predictability, on the other hand, is more stringent. In the context of emergence it would mean that one possessed a theory such that once one had an association between a base and an emergent, one could predict how a change in the base would modulate the emergent itself. In other words, a theory would be needed which could predict which emergent would be around in the situation of any given base.

³ For example physicist and reductionist Steven Weinberg writes: "Although it would be hard to explain the properties of a tornado in terms of the physics of electrons and quarks, I see this as a matter of calculational impasse, not an indicator of the need for new physical laws" (Weinberg (1993), p. 17).

4 which kind of emergence is interesting? evaluation of the framework

We are now in a position to give an overview of different kinds of emergence. We categorise these different notions along a single axis, from *trivial* to *interesting* to *incomprehensible*. This axis reflects primarily philosophical interests and secondarily scientific ones. By 'philosophical interests' we mean the position that emergence is interesting if it can challenge reductionism and clarify the notion of nonreductive materialism. We also show the relation of each notion of emergence that we present to the different kinds of reductionism, and provide examples when extant. The table is as follows:

	Notion of emergence	Comment	Relation to reductionism	Example
trivial	Pure phenomenological emergence	Phenomenological emergence consists in emergent properties or behaviours which are at first sight surprising for the observer, but after a closer look at the lower level their appearance is explainable and no longer surprising. The emergent character is due to our perceptual limitations.	Compatible with any form of reductionism	Chaotic attractors
	Epistemic emergence	Epistemic emergence consists mainly in properties or behaviours that appear on the higher level but are reducible in the sense of Nagel-reductionism. The reason for the existence of a theory for the upper level is basically an instrumental one, as the description of the phenomena is more compressed using the upper level theory.	Compatible with any form of reductionism, except the strong form of methodological reductionism (which nobody really holds anyway)	Probably many examples of complex systems theory, like pattern in CA, and ABMs Thermodynamics
interesting	Emergence of macroproperties	The emergence of macroproperties, of structural or functional organisation in a self-organised process, is an interesting case. First, it is not clear if the knowledge concerning the organisation belongs to the lower or to the upper level. Second, the process generates boundary conditions which influence the future progress of self-organisation. Generalisations concerning the forming of these boundary conditions are perhaps also not part of the theory of the lower level.	Not compatible with methodological reductionism; probably not compatible with epistemic reductionism	BZ-reaction

	Notion of emergence	Comment	Relation to reductionism	Example
interesting	Theoretical emergence	Theoretical emergence concerns primarily laws. Some laws appear on a upper level because their applications need a certain minimal degree of structure/organisation. Evolutionary processes for example need physically represented information which can mutate. The question in this cases is, if such laws have the same status like the laws on the basic level.	Not compatible with methodological and epistemic reductionism, probably not compatible with explanatory reductionism	Evolutionary theory
	Weak causal emergence	Weak causal emergents are properties which influence the lower level causally in a way that the possible causal space of the parts (the possible actions the parts in principle could cause given the laws of the lower level) is not violated.	Not compatible with methodological and epistemic reductionism, probably not compatible with explanatory reductionism	Social systems
incomprehensible	Strong causal emergence	Strong causal emergents are properties which change the possible causal space of the parts of the lower level.	not compatible with any form of reductionism	telekinesis?
	Mystic emergence	A mystic emergent property, law or macroproperty appears at a certain level and there is no way to really understand it. It has to be accepted as a primitive component of nature.	not compatible with any form of reductionism.	vitalism creationism (?)

table 3: framework

5 emergence and the science of complexity

Now that we have categorised concepts of emergence into the *interesting*, the *semi-interesting*, and the *uninteresting*, it is time to consider particular phenomena discussed in the complexity science literature, to see which of these emergence concepts they fall under. We will look at three venues in which emergence is postulated: cellular automata (CA), agent based models (ABMs), and the information-based approach of Crutchfield and Shalizi. After considering these particular fields, we will summarise the results of a survey of students and researchers at the Santa Fe Institute (SFI) about the concepts of emergence and reduction.

5.1. Cellular Automata Many complexity scientists consider patterns generated by some cellular automata (CA), those 'on the border between order and chaos', (a.k.a. Wolfram's class 4 rules which evolve complex localized structures), to be emergent. There are two issues to be considered. First, what is it about these particular patterns which leads researchers to say that they are emergent? Are the requirements based on necessary and sufficient conditions, or based on prototypes? Second, what type of 'emergence' is this, and how does it fit into our framework for emergence already discussed?

First of all, class 4 CAs, those most often claimed as exhibiting emergence, do not seem to have strict membership requirements (Mitchell (1998)). This gestures towards one of two situations: one, it

might be that our understanding of CAs is presently limited; while CA emergence has rigorous conditions, we don't yet understand them. On the other hand, it could indicate that there *aren't* rigorous conditions *at all*, and our concept of emergence in CAs, rather than admitting necessary and sufficient conditions, is instead defined by prototypes. According to this theory, the concept, a cognitive entity, is composed of what we consider to be good examples of emergence, coupled with a similarity measure which we use to group new putative emergence cases into or out of the category *emergent*. We are unable to decide which of these options is the case at this time – both are consistent with what we know about emergence in complexity science. Instead, we will consider some of the criteria, albeit imperfect ones, which have been proposed as ruling well on which CAs exhibit emergence.

Researchers provide a variety of justifications for the claim that patterns in cellular automata are emergent. First of all, some claim that there are connections with various ideas in computation theory. For example, Wolfram has claimed that class 4 CAs are capable of universal computation. However, as Mitchell has pointed out, this hypothesis is impossible to verify, since "there can be no general method for proving that a given rule is or is not capable of universal computation"(Mitchell (1998)). Langton also thought that he had a rubric for class 4 CAs, based on his lambda statistic, which for a binary CA, is the fraction of states (of 8) with transitions to the next time point in which the cell is "on". He thought that complex behavior in CA's could be correlated with the lambda parameter. However, there is scepticism about whether this criterion makes a good match for intermediate lambda values. Some claim that higher-level structures, like gliders, are emergent because it would be difficult to either *derive* or to *predict* their properties from the lower level rules which define the CA. Holland says, in particular, that "no extant analytical technique will predict the existence of a glider pattern. We can only discover the glider by observation, watching the laws play themselves out in different configurations"(Emergence 140). Notice the ambivalence („no *extant* analytical technique") about whether these higher-level properties are in fact impossible to predict, rather than merely not predictable at the present time. This ambiguity permeates writing in complexity science – few seem confident about whether predictions are impossible, or merely very difficult. The same holds for derivability.

Researchers analyzing the ability of evolved CAs to do computations have also found it useful to describe CA patterns at a level higher than the lower level rules which we usually see as defining CAs. Hordijk, Crutchfield and Mitchell (1996) describe how a CA is capable of determining whether the majority of cells in the initial condition in a 1D CA are „on" by redescribing the CA world in terms of the „laws of ,particle physics"". That is, they consider that, particularly with a CA whose rules were evolved, it can be difficult to understand how the computations are done by studying the basic CA transition rules alone. But, they find, such comprehension is more natural when the CA is redescribed in terms of higher level particle physics description (aka computational mechanics), for which they have a model. The claim might be either that the CA's computations are intrinsically better described on this other level, or that this is just a convenient way for us to understand what is going on. If it is a matter of scientific convenience, this might be promoting an epistemic view of emergence. If it is a matter of necessity, with the claim that the computational strategy is really only describable on the particle level, this could be a claim of the emergence of macroproperties.

In addition to the aforementioned criteria, some claim that CA emergents are so because they are *surprising*, independent of their predictability or derivability. In the cases of both CAs and ABMs, there is some waffling about the importance of 'unexpectedness' or 'novelty' and 'surprise' in the definition of emergence. Holland claims that, while we often say that unexpectedness is a hallmark of emergence, the unexpectedness of something is merely a *heuristic* that we use to guide us in looking for where emergence might be found. He explains as follows: "It is tempting to take the inability to anticipate—surprise—as a critical aspect of emergence. It is true that surprise, occasioned by the antics of a rule-based system, is often a useful psychological guide, directing attention to emergent phenomena. However, I do not look upon surprise as an essential element in staking out the territory. In short, I do not think emergence is an 'eye-of-the-beholder' phenomenon that goes away once it is understood"(Holland (1998), p. 5). Holland's view of how surprise works into to emergence, in CA's and elsewhere, makes sense in the light of how Complexity Science carries out investigations. While people often mention that some entity is unexpected, they rarely say that the entity ceases to be emergent once it is understood, which is what would happen if surprise, a phenomenological definition of emergence, were being exploited. Nevertheless, some people do include an intrinsically

subjective element in their view of emergence, in CAs and elsewhere. For example, in his textbook on complex systems, Bar-Yam (1992) says that behavior is emergent if „the behavior of the system cannot be simply inferred from the behavior of the component“. Until he provides a definition of ‚simply inferred‘ which isn’t merely linked to how a human will tend to think, it seems that a phenomenological theory of emergence is being supported.

Considering the reasons presented by complexity scientists for the emergence of some CA patterns, two conclusions are evident. First, there are a variety of criteria used to pick out emergents, and they are not monolithically applied. Second, it doesn’t seem that emergence in CAs, as seen by complexity scientists, is in opposition to reductionism. In fact, it often seems that analytical techniques *require* reductionism. Holland concurs: “The program for studying emergence set forth here depends on *reduction*”(Emergence 13). CA research seems to be dependent on reductionism in terms of explanation. When an interesting phenomenon is discovered, it is often explained by considering how the rules which define the CA could lead to the putative emergent entities. True, often the connections between the emergents and the lower level seem unnatural or surprising, but that doesn’t change the methodological task of the reduction. But note that, while supportive of the reductionist idea, pro-reductionist emergentists still distinguish between *simple* reduction, which they reject, and more moderate reduction, which is embraced. Simple reduction is reduction in which you attempt to explain the behaviors of an entire system by breaking down the system into parts, and then studying those part in isolation. To determine the behavior of a system with many parts, a simple reductionist would merely sum the behaviors of the parts observed when they act when alone. While often used as a reductionist straw-man, it isn’t clear that anyone actually hold this hyper-reductionists view. On the other hand, a moderate reductionists would admit that there might be non-linear interactions between the reduced parts, and that the explanatory projects would not be complete until a system is reduced to a description of the parts coupled with a description of the rules which govern interactions between the parts, rules which might include non-linear elements. This type of reductionism is still consistent with even the most hard-core of reductionists, like Weinberg in his chapter called „Two Cheers for Reductionism“(Weinberg (1993)).

5.2. Agent-Based Modelling Behaviours of ABMs are often considered to be emergent. As with CAs, our task here is to determine why researchers make such claims of emergence in ABMs, and then to determine which type of emergence outlined in section 4 is being called upon. Furthermore, we will consider whether the methods of ABMs are compatible with reductionism.

As with CAs, ABMs are often constructed from the bottom up – rules governing the interactions between parts are specified, and then the system composed of many parts is allowed to run, with the experimenter monitoring the results. Because they are models, some scientists have decided to take a design-perspective on emergence in ABMs which allows them to accentuate contrasts between agent based models and more traditional ones. For example, Clark (in Boden (1996)) suggests that emergent behavior is that which is „not neatly attributable to any specific inner component or system“. So to determine emergence here, one needs to know the control system in question and have an unequivocal description of it, as one would in the case of ABMs, but not necessarily so in the case of a real world phenomenon like *life* or *social behaviors*. Clark also requires that the behavior be only describable in a language with different predicates than one suitable to describe the agents themselves. For something not be emergent, the property would have to be influenced by a controlled variable, which is one that tracks „ behavior or properties which can be directly affected“.

As an example of an ABM which seems to display emergence according to this definition, consider Resnick’s termite wood-piling model. In this model, termites are agents living on a plane following the following two rules: 1) if encounter a woodchip and empty-handed, pick it up, 2) if carrying a woodchip and encounter another, put down the one being carried. If you make the termites turn at random and change direction, and you let this run, a distributed group of wood chips will end up piled into larger piles. Nevertheless, there is no ‚controlled variable‘ in this system which indicates that the piles should be condensed, since it is just as easy to take a chip off a pile as to add one to a pile. Yet, in fact, after 20,000 iterations, 2,000 scattered chips are condensed into large piles (Clark (1996)). To some, this indicates emergent behavior in ABMs, because there is no ‚wood-chip piling‘ parameter.

Another way to look at emergence in this case is by considering whether ‚more comes out than was put in‘. For example, Holland points to the unexpectedness of a result, the fact that the designers

didn't anticipate the result, as evidence that it is emergent: "The unanticipated predictions ... provide some of our best examples of emergence. A great deal more comes out than the authors anticipated, even allowing for their superb intuition"(Emergence 12). In the case of Resnick's model, this could mean that the designer (Resnick) wasn't expecting that the termites, as a community, would behave as they did, and their surpassing of the designer's expectations defined those behaviors as emergent. This take on emergence in ABMs makes it more phenomenological – the emergence is based on the anticipatory ability of the designer of the system. As we have mentioned, phenomenological emergence is perfectly consistent with reductionism, and seems unnecessarily subjective, since a not-so-bright designer will end up with many emergent entities, while a smarter or more experienced one would have none.

Discussions of both ABMs and CAs reveal some differences of opinion. In particular, the commitment to phenomenological emergence is varied. However, there is some common way of looking at emergence and reduction in these systems as follows. Micro-laws are seen as determining behavior, while macro-laws are seen as describing behavior. Often in science, we notice macro-level regularities, like gliders or woodchip piling behaviors. Investigation proceeds by looking for the lower-level laws which, if in force, would lead to a world in which the higher-level regularities existed. This is reduction or some sort. In the case of ABMs and CAs, we know that such lower-level laws exist since they were programmed as part of the system. So, when we observe an emergent, the explanatory project is both reductionist (how did we get this higher level behaviors out of some lower-level rules?) and guaranteed some success because the system is generated from lower level rules. So, although it seems that in these model systems, which are popular methods for investigating complex systems, reduction may be possible, the import of this for putative emergents in the natural world remains in question, since in that case, it is a matter of reductionist faith that there are lower level rules which generate the higher-level regularities.

5.3 the Crutchfield-Shalizi-approach to emergence James P. Crutchfield and Cosma Shalizi (who provided a clarification of Crutchfield's approach) are among those complexity scientists which claim to have formalised the notion of emergence (Crutchfield (1994a/b), Shalizi (2001)) (another attempt has been undertaken by Baas (1994)). We will here present this approach: Crutchfield differentiates between three aspect of emergence: intuitive emergence, pattern formation and intrinsic emergence. The notion of intuitive emergence captures the aspect of "newness" of an emergent entity in the sense that it could not have been expected by the observer of a system which displays the emergent entity. The notion of pattern formation goes a step further in clarifying the ontological aspect of the emergent entity. Emergent is therefore an organisational or structural aspect of the system, which appears in the temporal development of the system. The examples quoted by Crutchfield are well-known: deterministic chaos (unpredictability as emergent entity), self-avoiding random walk (self-similarity as emergent entity), Bénard-cells in heat-convection or patterns in the Belousov-Zhabotinsky reaction. But pattern formation still is insufficient for defining emergence in a scientific meaningful way, because of the crucial role of the observer. The observer actually has to recognise the pattern as such – and biases of the observer could prevent this recognition. The historical denials of patterns in the BZ-reaction serve as an example of this problem. Crutchfield therefore introduces the notion of intrinsic emergence: Emergent entities (which Crutchfield calls 'patterns') are those which the system itself creates in its evolution and which the system can use or capitalise on. The observer is somehow in the system itself. This idea is developed in the context of a computational view of nature: All we can know about a system are discrete time and discrete space measurement series. This string of symbols can be analysed using the tools of computation theory. Crutchfield developed the theory of computational mechanics [Crutchfield] which he claims is the appropriate tool to reconstruct the causal structure which underlines the generation of the measurement string. This causal structure is captured by so-called epsilon-machines, which have been proven to be the minimal description of this type. In this framework, the notion of emergence can now be formalised: roughly, a process undergoes emergence if at some time the architecture of information processing has changed in such a way that a distinct and more powerful level of intrinsic computation has appeared that was not present in earlier conditions.

Shalizi presented a further clarification of this concept of emergence: First, he rejects a notion of emergence which is a contraposition to explanatory reductionism: an entity should be called emergent if it cannot be explained from the properties and interactions of the parts of the lower level. Shalizi's point against this concept of emergence is that it would be impossible to know if anything is emergent

in this sense, because one could always claim that we just don't have at this time the explanation but we might have it in future. He rather proposes a much moderate version of emergence, based on epistemic emergence: emergent entities are those which arise from the interaction of the parts of the lower level but which the latter themselves do not display – temperature serves as a standard example. This idea is then combined with the prediction of a process. We will sketch the idea while referring the details to the original publication. This process relies on the causal structure of the system and is the source of the measurement-values, from which we gain our knowledge about the system. Understanding the system means reconstruction of the epsilon machine which itself can be used to predict the systems behaviour. This leads to a measure for efficiency of the prediction. The hierarchical structure (one of the assumptions for applying the concept of emergence) is imposed by introducing another process, which can be derived from the first by a function. If this new process has a greater predictive efficiency, it is called emergent. In the context of the epsilon machine this would lead to a partitioning of the machine in sub-machines, which are somehow stronger connected than others (e.g. loops). This notion of emergence – compatible with reduction – is also claimed to be observer independent.

How do we judge the Crutchfield-Shalizi-approach using our framework? Both authors point out the difficulty identifying which entities should be called 'emergent'. Crutchfield finally ends up with the notion of "pattern," with patterns the result of the changing causal structure of the process. Compared to our framework, the important entities are therefore the macro-conditions of the system – but formulated in a computational framework, as the structural arrangements of the causal states. A change in this arrangement leads to intrinsic emergence. In Shalizi's terminology, the emergent entities are processes, which are sequences of random variables, which arise out of the analysis of the obtained measurement-values of the system. In contrast to Crutchfield, Shalizi does not refer to the evolutionary change of the system which leads to a new causal structure which gives rise to emergence. He rather focuses on two different levels of descriptions: One kind of measurement leading to process S , another kind of measurement leading to S' and if there is a function which transforms S into S' and if S' has more predictive efficiency, then S' is emergent. As Shalizi notes, there might be a problem concerning observer-independence. There is still the act of choosing a new set of variables leading to S' . It cannot be excluded that there are different ways to choose S' and therefore different emergent properties arise. One way to choose between them might be to compare the predictive efficiency of the two choices. But we have to state, that Shalizi does not clarify Crutchfield, as he refers to a different aspect of emergence: Crutchfield's notion refers to the aspect of evolutionary changes of a system which will be reflected in the causal structure of the system. Shalizi refers to different levels of description of a system at a certain time of the system's history. Shalizi's concept can better be compared to our framework, where it turns out that he stands for a concept of emergence compatible with reductionism, whereas in Crutchfield's concept it is still somehow undefined, which the two levels are.

However, both approaches are interesting, not only because of their notions of emergence, but because they make strong claims concerning what should count as an explanation. They say that the reconstructed causal states are the best model we can get of the system under investigation. But this conception seems, at least for us, to be in a strong opposition to what science generally counts as an explanation of a system: for example, chemical bonds are explained in terms of quantum theory, the metabolism in a cell is explained by a set of chemical reaction chains, the ontogenesis of *C. elegans* is explained by referring to a cascade of gene-activation, the planetary movements are explained by referring to Newton's laws, etc. – but nowhere do epsilon-machines appear. It is still a matter of debate if the computational view on nature is the appropriate one that fits into the generally accepted notion of scientific explanation. This might be the main question that needs to be answered in order to judge Crutchfield and Shalizi's approaches.

5.4 emergence and the SFI community – results from a non-representative survey

Students of the 2002 Complex Systems Summer School in Santa Fe and SFI researchers were questioned about emergence, prediction, reduction and causation by the authors via a 32-question email survey. In order to further determine the role of the emergence concept in complexity science, we have analysed these results. This population of researchers is a particularly good one to survey vis-à-vis emergence because both groups have both education and research interests in complexity science. However, because only 32/60 students, and 6/50 researchers responded, as well as the fact

that the survey was untested for comprehensibility, we only consider these results with substantial hesitation.

As with the ideas considered in the previous three sections, the concept of emergence seems to be used by researchers and students in a variety of ways, with no absolute consensus about meaning. Despite differences of opinion about the details of emergence, the importance of emergence was generally agreed upon: 93% of students and 67% of researchers reported in the affirmative that "the concept of emergence is important for the science(s) of complexity and it should be one of the goals of this science to enhance our understanding of emergence and to quantify somehow this notion". While in the clear minority, it is also interesting to note that 15% of students and 33% of researchers think that 'emergence' is a buzzword to be avoided, while 22% of students and 16% of researchers think that it is a filler term which should be replaced when an entity is understood. This last opinion, that an entity is emergent only until it is better understood, seems to be in-line with the phenomenological emergence mentioned in section 4, a variety of emergence seen as philosophically trivial because the emergence is observer-dependent.

We've discussed ontology, that is, which kinds of entities the term 'emergence' should apply to. We asked the SFI community what they thought about the ontology of emergence. The results suggest that many think that 'emergence' can apply promiscuously, to everything from laws to patterns to behaviours. Similarly, 'emergence' applies to many of the systems or methods often used in complexity science, said the SFI community. Over 60% of the two communities thought the following five systems were emergent: consciousness, self-organisation, agent-based models, attractors in ND systems, and cellular automata.

Given the connections between emergence and reducibility, we asked about the connections between different levels of description, and whether or not the higher-level emergents were reducible to the lower level. There was substantial disagreement on this topic. Among students, responses were almost evenly split between the following 3 options: 1) emergents are reducible now to the lower level, 2) emergents will in time be reducible to the lower-level, 3) emergents will never be reducible to the lower level. Researcher results were similar, although with more thinking that emergents were presently reducible. Note that, considering the emphasis in the complexity science literature that emergence is somehow *opposed to* reductionism, it is somewhat surprising that about 2/3 of those surveyed thought, it seems, that there isn't tension between emergence and reductionism, in that emergents are themselves, or will be soon, reducible to the lower level.

Connected with reduction, we inquired about the presence or absence of downward causation in emergence. The majority (62% of researchers and 75% of students) thought that 'downward causation' was necessary for emergence, although the downward causation which they seemed to support was not the same as the downward causation so loathed by philosophers for its logical inconsistency. The survey participants seemed to think that "downward causation is basically the result of the structural or functional organisation of the parts on the lower lever (e.g. a feedback mechanism)". This sort of downward causation does not allow for the higher-level emergents to break lower level laws, and seems most in line with the emergence of macroproperties described in section 4.

In the end, the survey indicated some differences of opinion about the meaning and applicability of emergence. However, the take home message seems to be that most think that emergence is compatible with reductionism, evidenced by the fact that they both ruled-out rule-breaking downward causation and by think that reductions either have, or will soon, be provided by scientists.

6 conclusion

Our investigations have lead us to the following conclusions:

- 1) Many of the emergent entities complexity science tries to explain fall into the aforementioned category of epistemic emergence. Recall that epistemic emergence, although countenancing the existence of higher-level 'emergent' entities, still claims that they are reducible to lower-level descriptions. In the case of CAs, although researchers find it convenient to look at gliders and other 'higher-level' phenomena as emergent, they don't deny that they can be perfectly well

described, and explained, by the lower level rules. In fact, explanations of glider behaviour *depend* on the ability to reduce them to the lower-level. The same can hold for emergents found through agent-based models.

- 2) Our review, both of the literature and of the SFI community, has indicated that there is still some confusion concerning the use of the emergence concept. First of all, the existence of a certain deductive relationship between the lower and higher levels in science remains in question, even among complexity science researchers. The best way to prove that it is possible to deduce higher-level rules from lower level ones would be to provide lots of examples. Yet those examples have not been forthcoming – researchers claim, as a matter of faith, that such deductions, although possible, are just too complicated to feasibly create. Hence, there is a lack of clarity about whether the reductions are merely incomplete, or whether they are technically impossible to achieve. The role of prediction and novelty are also matters of confusion. Some claim that novelty is just a heuristic which can indicate to researchers candidate emergent properties, which are actually defined using other criteria. Others claim that the unexpectedness of emergents are a necessary condition for their existences. Pioneers in emergence in biology, Goodwin and Sole, claim that the question of emergence is tantamount to the question: "How can systems made up of components whose properties we understand well give rise to phenomena that are quite unexpected?"(Goodwin and Sole (2000)).
- 3) Based on our evaluation of the framework, we think that there exist interesting and coherent concepts of emergence from a philosophical point-of-view: emergence of macroproperties, theoretical emergence and weak causal emergence. Each of these three concepts are in conflict, or at least tension, with at least one of the reductionist claims. Yet these don't seem to be the concepts of emergence used by complexity science researchers anyhow.
- 4) Downward causation – seen by philosophers as deeply problematic – seems to be more coherent than is sometimes assumed. Weak causal emergence, which differs from strong causal emergence in that it doesn't require *violation* of lower level laws, seems to be a perfectly acceptable viewpoint. For example, note that self-organising systems create boundary conditions which subsequently change the behaviour of the parts at the lower level without changing the possible causal space.
- 5) There exists the possibility of showing an important aspect of non-reductionist materialism: Whenever scientists deal with a so-called "upper-level" problem, the first approach will be a non-reductive one: the scientist has to gain an understanding of the problem and he or she will do so not only in referring to a lower level theory, but to get the phenomenology of the upper level and create first generalisations. If they are somehow to be reduced to a lower-level theory happens much later – if ever – in the process of scientific understanding- So one main point of non-reductionist materialism might be an instrumental one.
- 6) While complexity science might be compatible with, or even supported by, reductionism, it does point out some interesting questions about the nature of that reduction which mainstream research programs miss; complexity science must confront head on the relation between the lower and upper level, and decide, for any higher level, which other level of magnification would be the appropriate one on which to ground the reduction and explanation. As there might be no problem with the assumption of hierarchical realism, the main problem is to identify the hierarchies (or levels or organisation) and the correct order!
- 7) We assume that the research program of complexity science is a reductionist one (although not a simple reductionist approach, see section 5.2). In that case, why is it that complexity researchers seem so committed to non-reductive materialism, and hostile to reductionist rhetoric? One explanation is that reductionism is considered to be naïve, and also to make the world seem to systematic and boring, eliminating all mystery, novelty, and excitement. Non-reductive materialism might then have some public relation value over reductionism.

works cited

- Boden, Margret (1996). The Philosophy of Artificial Life. Oxford U Press.
- Baas, Nils A. (1994); Emergence, Hierarchies, and Hyperstructures; In: Langton, Christopher G. (ed.) (1994); Artificial Life III; Addison-Wesley; Redwood City.
- Bar-Yam, Yaneer (1992). Dynamics of Complex Systems. Addison-Wesley, Massachusetts.
- Beckermann, Ansgar / Flohr, Hans / Kim, Jaegwon (eds.) (1992); Emergence or Reduction. Essays on the Prospects of Nonreductive Physicalism; Walter de Gruyter; Berlin, New York.
- Broad, C. D. (1925), The Mind and its Place in Nature, Routledge & Kegan Paul, London 1968.
- Casti, John L. (1986); On System Complexity: Identification, Measurement, and Management; In: Casti, John L. / Karlqvist, Anders (eds.) (1986); Complexity, Language, and Life: Mathematical Approaches; Springer-Verlag; Berlin.
- James P. Crutchfield (1994a). Is Anything ever New? Considering Emergence. In: G. Cowan, D. Pines, D. Melzner, Integrative Themes, Santa Fe Institute studies in the Sciences of Complexity XIX, Addison-Wesley, Reading MA.
- James P. Crutchfield (1994b). The Calculi of emergence. Computation, Dynamics and Induction, Physica D, special issue on the eProceedings of the Oji International Seminar Complex Systems – from Complex Dynamics to Artificial Reality.
- Goodwin and Sole (2000). Signs of Life: How Complexity Pervades Biology. Basic Books: Perseus Books Group.
- Holland, John H. (1998), Emergence. From Chaos to Order, Helix Books , Reading, Massachusetts.
- Hordijk, Crutchfield, and Mitchell (1996), Embedded-Particle Computation in Evolved Cellular Automata. In: Toffoli et al eds., Physics and Computation, '96 (Pre-proceedings), New England Complex Systems Institute, 153-158.
- Hoyningen-Huene, Paul (1994); Zu Emergenz, Mikro- und Makrodetermination; In: Lübbe, Weyma (ed.) (1994); Kausalität und Zurechnung. Über Verantwortung in komplexen kulturellen Prozessen; Walter de Gruyter; Berlin, New York.
- Kim, Jaegwon (1999), Making sense of emergence, Philosophical Studies 95, 3-36.
- Lewin, Roger (1999) Complexity: Life at the Edge of Chaos. University of Chicago Press.
- McLaughlin, B.P. (1992); The Rise and Fall of British Emergentism. In: Beckermann et al. (1992).
- Mitchell, M (1998); Computation in Cellular Automata; In: Gramss et al. Nonstandard Computation, Weinheim: VCH Verlagsgesellschaft
- Nagel, Ernest (1961/1979); The Structure of Science. Problems in the Logic of Scientific Explanation; Routledge & Kegan Paul; London, Henley.
- Oppenheim, P. / Putnam, H. (1958); The Unity of Science as a Working Hypothesis; In: Feigl, H. / Scriven, M. / Maxwell, G. (eds.) (1958); Minnesota Studies in the Philosophy of Science, vol. II; University of Minnesota Press; Minneapolis.
- Shalizi, Cosma (2001). Causal Architecture, Complexity and Self-Organisation in Time Series and Cellular Automata, PhD dissertation, University of Wisconsin at Madison, Physics Department.

Simon, Herbert A. (1962); The Architecture of Complexity; Proceedings of the American Philosophical Society 106, No. 6; p. 467-482.

Stöckler, Manfred (1991); Reductionism and the New Theories of Self-Organisation; In: Schurz, Gerhard / Dorn, Georg J.W. (Eds.) (1991); Advances in Scientific Philosophy; Editions Rodopi B.V.; Amsterdam - Atlanta GA.

Waldrop, M. Mitchell (1992); Complexity. The Emerging Science at the Edge of Order and Chaos; Touchstone, New York.

Weinberg, Steven (1993); Dreams of a Final Theory; Pantheon Books, New York.